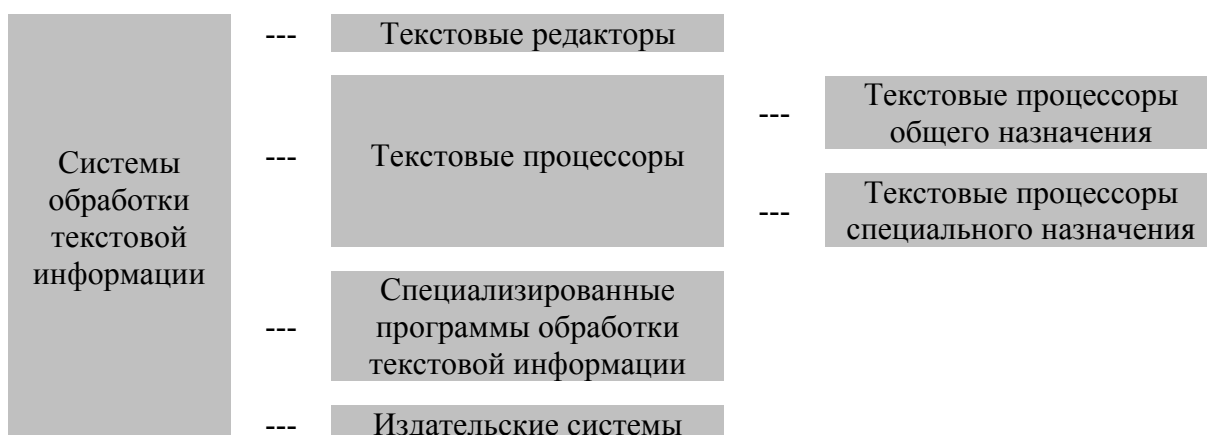


КЛАССИФИКАЦИИ СРЕДСТВ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ



Текстовый редактор - это инструментальное программное средство, предназначенное для создания и редактирования текстов, не содержащих сложных структур (параграфов, глав и др.), и имеющее малое количество функций.

К текстовым редакторам следует отнести **редакторы текстов программ** (встроены в языки программирования и поддерживают их синтаксис) и собственно **редакторы текстов**.

Основные функции:

1. Просмотр и редактирование простейших текстов (ReadMe файлы, описания и т.д.).
2. Редактирования исходных текстов программ.
3. Редактирование конфигурационных и настроечных файлов Windows.
4. Работа с HTML – файлами (источниками).

Тестовый файл может использовать ряд тестовых кодировок.

Каждый символ кодируется последовательностью битов. Обычно на **каждый символ отводится байт, т.е. 8 битов**. Однако же биты и байты - вещи довольно абстрактные и трактовать их можно по-разному. **Кодировка** представляет собой таблицу символов, где каждой букве алфавита (а также цифрам и специальным знакам) присвоен свой **уникальный номер - код символа**.

Стандартизирована только половина таблицы, т.н. ASCII-код - первые 128 символов, которые включают в себя буквы латинского алфавита. И с ними никогда не бывает проблем. Вторая же половина таблицы (а всего в ней 256 символов - по количеству состояний, который может принять один байт) отдана под национальные символы, и в каждой стране эта часть различна.

ASCII (American Standard Code for Information Interchange - Американский стандартный код для обмена информацией). Напомним, что этот формат оперирует с 256 численными кодами, имеющими значения от 0 до 255. В соответствии каждому коду ставится определенный символ (буква, цифра, знак препинания, математический символ или символ псевдографики). Это соответствие задается с помощью стандартных кодовых таблиц с различными номерами (например, таблица 866 предназначена для русскоязычных пользователей). Не содержит форматирования текста, поэтому является переносимым между различными операционными системами и программами.

ANSI (American National Standard Interface) – кодировка в среде Windows. У этих кодировок совпадают те части, которые относятся к латинскому алфавиту, специальным символам, цифрам, знакам препинания и математическим операциям, а различаются — относящиеся к другим алфавитам и псевдографике

В России умудрились придумать целых **5 различных кодировок**:

KOI-8 – возникла первой, для Unix-систем;

DOS-кодировка (или 866 кодовая страница). В ней были введены спецсимволы для рисования рамок (символы псевдографики);

Win-кодировка (или кодовая страница 1251);

ISO-8859-5 - ISO - международная организация по стандартам.

Сейчас идет работа над созданием новой универсальной кодировки (UNICODE), в которой предполагается в одну кодовую таблицу записать все языки мира. Тогда точно проблем не будет. Для этого на каждый символ отвели 2 байта. Таким образом, максимальное количество знаков в таблице расширилось до 65535.

Примеры:

Edit – редактор DOS

Встроенные редакторы программ-оболочек для DOS (Norton Commander, Dos Navigator и др.)

Блокнот (Notepad) – стандартная программа Windows. Размер файла ограничен 64 Кб. Notepad+, EditPad, Bred, Pad, UltraEdit, EditPlus.

Текстовый процессор - это инструментальное программное средство, предназначенное для создания и редактирования текстов сложной структуры и имеющее широкое функциональное наполнение. Чаще всего различают текстовые процессоры общего и специального назначения. Отличительной особенностью текстовых процессоров является возможность обрабатывать такие регулярные структуры документа, как абзац, параграф, страница и др. В мире существует сотни текстовых процессоров, различных по своему функциональному наполнению. Чем больше функций реализует тот или иной процессор, тем он более сложен для освоения. Среди наиболее распространенных - текстовые процессоры общего назначения: Multe-Edit, Xy Write, Microsoft Word, Word Perfect. Проводя сравнительную характеристику текстовых процессоров, оценивают, как правило, следующие их функции: редактирование, форматирование, слияние файлов, настольное издательство, печать и такую характеристику, как быстродействие.

Рассматривая функцию редактирования, особое внимание уделялось таким возможностям, как: максимальное количество окон, наличие команды Undo (откатки); поиск опечаток; использование перекрестных ссылок; обработка структурированных текстов; режим, редактор и язык программирования макрокоманд.

Функция форматирования оценивается в зависимости от наличия следующих возможностей: контроль за висячими заголовками; управление заголовками, многоколонковый набор, задание таблицы стиля.

Под слиянием файлов подразумевается возможность считывания файлов, подготовленных в системах Lotus 1-2-3, dBASE: считывание ASCII- файлов; возможность использования условного оператора и математики; создание табличной формы.

Важнейшими возможностями настольного издательства, которые чаще всего реализованы в текстовых процессорах - это импорт графики, предварительный просмотр страниц, размещение текста вокруг графики, рисование линий или прямоугольников с текстом.

При оценке функции печати рассматривают возможность организации пропорциональной разрядки, постановки документов в очередь на печать, фоновой печати и поддержки языка PostScript.

В нашей стране широкое распространение получил текстовый процессор Лексикон, который реализует такие функции, как: работа с абзацем, фрагментом, страницей; поиск и замена, форматирование, выход в DOS и многие другие.

Среди текстовых процессоров специального назначения следует отметить такие, как Unv Editor (до версии 4,0 этот процессор назывался ChiWriter) и Rt-chk, пригодных для

подготовки научных текстов, содержащих математические, физические или химические формулы, допускающие возможность использования до 20-ти различных шрифтов одновременно (для одного документа). Кроме того, они позволяют готовить документы с использованием верхних и нижних индексов, готическими, греческими, латинскими и русскими буквами и специальными знаками, такими как, например, или . Rt-chk имеет встроенный перекодировщик, работающий с 4-мя кодовыми таблицами (Основной, Альтернативной, КОИ-8 и Болгарской). Повысить эффективность и возможности текстового процессора позволяют специализированные программы коррективы документов.

Специализированные программы обработки текстовой информации - это программные средства, имеющие узкую специализацию. Среди таких программ следует отметить программы **проверки правописания и подбора синонимов, формирования текстов, перекодировщики, программы групповой записи текстов, словарные программы**. Многие текстовые процессоры в качестве функций содержат некоторые из таких программ. Однако, как правило, возможности встроенных программ ограничены по сравнению со специализированными. В качестве примера рассмотрим возможность орфографического контроля текста (иногда такие программы называют спеллерами - от английского слова speller).

Текстовый процессор Лексикон реализует функцию орфографического контроля. Однако словарь последней некоммерческой версии 6,67 содержит около 15 тысяч слов. Самой популярной специализированной программой по **проверке правописания** русских текстов является программа ОРФО. Словарь этой программы (V3.0) содержит 220 тыс. лексем (около 3.5 миллионов форм слов). Программа обладает уникальным словарем синонимов в 30 тыс. слов и выражений. В ОРФО реализованы такие функции, как поиск ошибок согласования слов в предложении по роду, числу и падежу; поиск нарушения правил пунктуации и неправильного использования заглавных и строчных букв; поддержка работы манипулятора "мышь". Программа является резидентной и совместимой со многими текстовыми процессорами (в том числе с Microsoft Word, Word Perfect, Multi-Edit, Лексикон).

Программы **формирования текстов** помогают составить (задать) общий вид документа. Кроме того многие из них могут сжать или расширить весь текст или его часть, а также, что делает документ более читаемым; при необходимости осуществить динамическое копирование заголовков (при сжатии или расширении одного из них соответственно изменяются и все другие). Некоторые программы формирования текстов содержат образцы форм текста; глоссарий; спеллер; поддерживают сортировку заголовков и сбор примечаний под одним заголовком. Программы могут загружаться в компьютер резидентно, автономно или совместно с текстовым процессором, базой данных или электронными таблицами. Среди наиболее распространенных: Grand Viwe (Symantek), Max-89, PC-Outline (Broun Bag Software).

Проблема совместимости текстовых процессоров в некоторых случаях может быть решена с помощью так называемых **перекодировщиков текста** или программ преобразования файлов. Среди наиболее распространенных - перекодировщик CONV, позволяющий достичь совместимости текстов, написанных в системах, использующих три различных вида кодовых таблиц (Основную, Альтернативную и Болгарскую). Другой перекодировщик с таким же названием работает с Альтернативной кодовой таблицей и старым вариантом КОИ-8. Тем самым он обеспечивает совместимость текстовых процессоров компьютеров IBM PC и Yamaha, последний из которых используется в некоторых высших педагогических учебных заведениях и школах (более подробно о перекодировщиках в следующих уроках). Кроме того существуют программы, обеспечивающие преобразование файлов, подготовленных на одном и том же типе техники, но с использованием различных процессоров текстов. Так, например, программа Software Bridge (V.3.11) обеспечивает преобразование форматов файлов

между 19 текстовыми процессорами (в том числе Word Perfect, Microsoft Word). Эта программа позволяет изменять подстрочные примечания, шрифты. А программа Word Transformer (V.4.01) преобразует файлы в формат ASCII, поддерживает 18 наиболее распространенных текстовых процессоров (в том числе Microsoft Word, Xy Write, Word Perfect).

Огромную помощь в подготовке документа, над которым одновременно работают несколько авторов, могут оказать программы **групповой записи текстов**. Они позволяют вносить поправки и комментарии в документ, не уничтожая оригинала. Некоторые из этой группы программ позволяют сравнить два текста и выделить в них, например цветом, отличные части. Другие - используют различные шрифты для изменений, вносимых в текст. Одной из наиболее популярных программ групповой записи текстов является программа ForComment, разработанная фирмой Braderbund. С документом могут работать до 15 соавторов. Внесения изменения помечаются именем соавтора и датируются. Однако вносить поправки в оригинал может только редактор. Программа позволяет хранить до 26 вариантов документа и до 15 замечаний или комментариев к каждой строке.

Словарные программы ориентированы на использование специалистами различных отраслей знаний и содержат определение слов и фраз. Многие из них могут: содержать перекрестные ссылки; отыскивать синонимы; давать множество определений для слов, имеющих более одного значения; проверять правильность написания слов. В качестве примера можно привести программу Stedmans Medical Dictionary, которая является медицинским словарем на 68 тыс. терминов. Программа Chase Wards позволяет дать определение 80 тыс. слов (в том числе и из "Краткого словаря по электронике издательства Webster") и 40 тыс. синонимов. Все эти программы позволяют наиболее профессионально редактировать различные документы.

Однако для подготовки такой печатной продукции, как книги, журналы, рекламные буклеты, в большей степени подходят **издательские системы**. Чаще всего они позволяют работать в режиме **WYSIWYG (What You See Is What You Get** - что Вы видите, то и получаете). Имея менее низкие показатели при наборе текста по сравнению со многими текстовыми процессорами, они располагают уникальными возможностями по их верстке. Среди наиболее типичных следующие возможности: загрузить текстовый или графический файл; осуществить элементарную корректуру текста; выделить участки текста иным шрифтом; осуществить многоколоночный набор; воспользоваться графическими примитивами (линия, прямоугольник, круг, окружность, закрашенный прямоугольник и др.); наложить объекты друг на друга; создать стилевой файл или новый шрифт; задать полиграфические параметры каждого абзаца и многое другое. Кроме того, издательские системы ориентированы на использование лазерных принтеров или фотонаборных автоматов. По общему мнению, самыми распространенными являются три программы верстки — **Quark XPress, Adobe InDesign и Adobe PageMaker**.

Верстая публикацию в одном из перечисленных пакетов, вы имеете дело с несколькими типами объектов. Это фреймы — контейнеры для импортированных извне текста и графики, и графические объекты, созданные в самой программе верстки. Кроме того, в PageMaker можно использовать еще один элемент — «простые» текстовые блоки. В PageMaker можно изменять размеры и положение текстовых и графических фреймов, но нельзя вкладывать один фрейм в другой, а в QuarkXPress можно вставлять только графические фреймы в текстовые. В этом отношении InDesign несравненно более удобен: он позволяет создавать очень сложные макеты с использованием множества фреймов, вложенных один в другой.

Неоспоримое преимущество InDesign — многократная отмена и повтор действий при верстке, в отличие от однократных отмен (да и то не всегда возможных) в XPress и PageMaker.

Одна из фундаментальных особенностей верстальных программ — использование шаблонных элементов для «сквозного» оформления публикаций. Служат для этого мастер-страницы, на которых и располагаются повторяющиеся шаблонные объекты.

Самым стабильным (особенно при работе с большими публикациями) и производительным пакетом является Quark XPress, и это одна из причин его многолетнего лидерства на рынке издательских программ. Он же и самый нетребовательный к ресурсам компьютера: работать с XPress 4.1 можно, используя старенькую «четверку» (не Pentium 4, конечно, а i486) с 12 Мбайт памяти.

PageMaker в этом отношении — твердый середнячок, а вот InDesign удивительно нетороплив и требует изрядных вычислительных ресурсов.

Программы подготовки электронных версий документов.

Разработанный фирмой Adobe формат Acrobat PDF стал стандартом для совместной работы с электронными документами, поскольку при его использовании вы можете быть абсолютно уверены, что отосланная вами страница будет воспроизведена в точно таком же виде. В программе реализованы инструментальные средства для коллективной работы над документами, обеспечения конфиденциальности и форматирования. Поэтому ее применение стало практически обязательным при передаче файлов через Интернет и выполнении определенных межплатформных работ, в том числе при публикации электронных книг.

В процессе инсталляции Acrobat размещает в вашей системе входящую в состав пакета программу Distiller, которая в фоновом режиме формирует PDF-файлы почти из любого документа. Аналогичным образом Acrobat помещает пиктограммы и выпадающее меню в MS Office для формирования PDF-файлов; в версии 5.0 значительно упрощено преобразование PDF-файлов в формат RTF и обратно. Созданные в Acrobat формы могут встраиваться в Web-страницы и присоединяться к серверным базам данных. С помощью Acrobat очень легко выделять графические объекты из PDF-файлов и использовать их в других документах. Даже Web-страницы (или целые узлы) могут быть загружены из Интернета в виде PDF-файлов - неограниченная возможность при проведении исследовательских работ, а также при архивировании.

ШРИФТЫ. КЛАССИФИКАЦИЯ. ХАРАКТЕРИСТИКИ

Шрифт (фонт)- это система отображения символов (текстовых, значков) на экране и при печати. Шрифт состоит из изображений (рисунков) отдельных символов - **глифов (glyph)**.

Для внутреннего представления глифа в файле шрифта True Type используются описания контуров, причем один глиф может содержать несколько контуров (рис. 1).

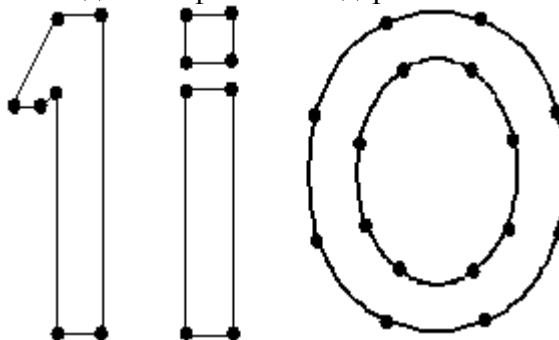


Рис. 1. Рисунки символов

Глифы могут иметь различный внешний вид (typeface). Операционная система Windows классифицирует шрифты на несколько типов, или семейств (font family). Эти типы называются Modern, Roman, Swiss, Script, Decorative.

Шрифты семейства Modern имеют одинаковую ширину букв. Шрифты семейства Roman содержат буквы различной ширины, имеющие засечки. Семейство Swiss отличается тем, что при переменной ширине букв они не имеют засечек. Буквы в шрифтах семейства Script как бы написаны от руки. Семейство Decorative содержит глифы в виде небольших картинок (пиктограмм).

Отдельные шрифты группируются в семейства сходных шрифтов, отличающихся друг от друга некоторыми характеристиками, такими, как стиль, размер, различные специальные эффекты.

Оперируя со шрифтами, часто пользуются следующими терминами:

- **размер шрифта или кегль.** Определяет высоту символов, обычно измеряемую в пунктах, составляющих 1/72 дюйма или примерно 0.36 мм, но в различных программах единицы его измерения можно устанавливать самостоятельно; 1 пункт = 0,376 мм – Россия и др. страны; 0,351 мм – США и Великобритания.
- **стиль шрифта (начертание).** Бывает четырех основных видов: обычный (normal), курсив (italic), полужирный (bold) и полужирный курсив (bold-italic). Есть и другие виды, например, утонченный (light), но они уже относятся к числу специальных, дизайнерских;
- **эффекты.** Включают используемый цвет, подчеркивание и другие дополнительные эффекты, например, перечеркивание, в некоторых случаях закрашивание и т. д.;
- **интервал.** Определяет расстояние между символами на экране или на печатной странице. Интервал между символами может быть как фиксированным (их мало), так и автоматически настраиваемым в зависимости от того, какие символы оказались расположенными рядом (так называемый кернинг);
- **ширина символов.** Оказывает воздействие на ширину всех символов данного шрифта.

По ширине шрифт может быть обычным, плотным или растянутым. Шрифты делятся на *моноширинные* и *пропорциональные*. Моноширинные шрифты, например, Courier, используют символы одинаковой ширины, а пропорциональные, например, Arial - различной, изменяющейся в зависимости от начертания конкретного символа. Кернинг – настройка пространства между индивидуальными знаками.

По области использования все шрифты делятся на наборные и декоративные «акцидентные».

Наборные гарнитуры (**Гарнитура - комплект шрифтов одного начертания.** Например, Arial, Arial Курсив, Arial Полужирный) - это рабочие лошадки: ими набирают основное тело текста любой длины. При установке Windows вы сразу получаете в свое распоряжение три таких гарнитуры: Times New Roman, Arial, Courier New.

Декоративные же шрифты предназначены для выделения текста, то есть для всевозможных заголовков, титулов, рекламных призывов др.

Стандарт Windows включает поддержку нескольких технологий **растровых** и **векторных** шрифтов.

Растровые шрифты. Хранятся на диске в виде набора битов, отображаемых на элементы экрана – пикселы. Создает на экране фиксированную картинку раstra, являющуюся организованным набором пикселов. Представляют собой точечные изображения, хорошо приспособленные для быстрого вывода на экран. Иными словами, это специальные, служебные шрифты, которые использует для своих нужд сам компьютер. В составе Windows поставляется пять базовых растровых шрифтов: MS Serif, MS Sans Serif (8, 10, 12, 14, 18, 24), Courier (10, 12, 15), Small Fonts (2, 4, 6) и Symbol (8, 10, 12, 14, 18, 24). Каждый кегль хранится в отдельном файле с типом *.fon.

Векторные шрифты. Шрифты этого типа можно масштабировать, то есть увеличивать или уменьшать их, умножая на любой коэффициент линейные размеры. Символы

хранятся в виде множества точек и соединяющих их линий, и эту информацию Windows может использовать для построения шрифта произвольного масштаба. Даже будучи увеличенными во много раз, векторные шрифты не теряют своего внешнего вида и остаются гладкими.

Векторный фонт < 18 пунктов плохо читаем на дисплее, >12 пунктов хорошо смотрится на принтере с разрешением св 300 dpi.

Область использования: вывод на плоттер или др. устройство черчения. Roman, Script, Modern.

Самыми распространенными среди векторных шрифтов являются так называемые шрифты True Type. В Windows технологии True Type отдается предпочтение по сравнению с другими шрифтовыми технологиями. Она была специально разработана Microsoft, чтобы предоставить стандартные шрифты, пригодные как для отображения на экране, так и для вывода на печать, и при этом обеспечить максимальное сходство получающихся результатов. Каждый фонт содержит набор рекомендаций – хинтов. В хинтах заключена информация как модифицировать форму знаков. Размер 1-999 пунктов (если нет ограничений в приложении).

В Windows каждый из шрифтов True Type хранится в файле с расширением TTF. Он содержит информацию о форме символов и коэффициенты, необходимые для масштабирования шрифта. В составе Windows поставляется большое количество шрифтов True Type, включая группы Arial, Courier New, и Times New Roman, каждая из которых содержит несколько шрифтов различных стилей (обычный, курсив, полужирный и полужирный курсив).

Нельзя не упомянуть шрифты PostScript. Такие шрифты хранятся в виде двух файлов каждый с расширениями PFB и PFM. Они используются в полиграфическом производстве, и для того, чтобы они корректно отображались на экране и распечатывались, на компьютер нужно устанавливать специальную дополнительную программу управления шрифтами.

В Windows существует специальная системная папка шрифтов WINDOWS\FONTS. DLL-библиотека commdlg.dll содержит функцию ChooseFont, специально предназначенную для выбора одного из зарегистрированных в системе шрифтов.